PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: METHODS FOR DETECTING CANCER-ASSOCIATED CHROMOSOME DELETIONS USING IN SITU HYBRIDIZATION

(57) Abstract

The present invention provides accurate and efficient means of detecting chromosomal deletions associated with LOH and neoplastic growth. In particular, it provides methods and compositions for the localization of precise regions of LOH on chromosome 16q in breast tumor cells, and identifying tumor suppressor genes encoded by this region.

METHODS FOR DETECTING CANCER-ASSOCIATED CHROMOSOME
DELETIONS
USING IN SITU HYBRIDIZATION

5

BACKGROUND OF THE INVENTION

The present invention provides accurate and
efficient means of detecting chromosomal deletions
associated with neoplastic growth. In particular, it
10   provides methods and composition for the localization
of precise regions of loss of heterozygosity on
chromosome 16q in breast tumor cells, and identifying
putative tumor suppressor gene(s) encoded by this
region.

15   Genetic aberrations in a single cell are
currently hypothesized to give rise to cancer (Nowell,
Science 194:23 (1976), Foulds, *J. Chronic Dis.*
8:2(1958). Identification of the genetic events
leading to neoplastic transformation and subsequent
20   progression can facilitate efforts to define the
biological basis for disease, improve prognostication
and prediction of therapeutic response, and permit
earlier tumor detection.

In particular, the deletion or multiplication
25   of copies of whole chromosomes or chromosomal segments,
and higher level amplifications of specific regions of
the genome are common occurrences in cancer. See, for
example, Smith, *et al.*, *Breast Cancer Res. Treat.*, 18:
Suppl. 1: 5-14 (1991), van de Vijer & Nusse, *Biochim.*
30   *Biophys. Acta.* 1072: 33-50 (1991), Sato, *et al.*,
Cancer. Res., 50: 7184-7189 (1990). In fact, the
amplification and deletion of DNA sequences containing
proto-oncogenes and tumor-suppressor genes,
respectively, are frequently characteristic of
35   tumorigenesis. Dutrillaux, *et al.*, *Cancer Genet.*
Cytogenet., 49: 203-217 (1990). Clearly the
identification of amplified and deleted regions and the
cloning of the genes involved is crucial both to the
study of tumorigenesis and to the development of cancer
40   diagnostics.

2

One model of carcinogenesis involves the
inactivation of tumor suppressor genes that, in
noncancerous cells, are believed to regulate cellular
proliferation, differentiation or morphogenesis.
5   Inactivation of these genes can lead to uncontrolled
cell proliferation, resulting in carcinogenesis.
(Hollingsworth et al., J. Natl Cancer Inst. 83, 91-96
(1991); Solomon et al., Science 254, 1153-1160
(1991)). Typically, carcinogenesis requires that both
10   alleles of a tumor suppressor gene be inactivated
(Knudson, Proc. Natl. Acad. Sci. USA 68, 820-823
(1971)). Often, inactivation occurs by mutation of one
allele and loss of the other as a consequence of either
mitotic recombination or physical deletion.
15   Consequently, tumors carrying mutant tumor suppressor
genes often show loss of heterozygosity (LOH) near the
site of the suppressor gene. Several tumor suppressor
genes, for example, RB1 (13q14.2), TP53 (17p13.1), WT1
(11p13), APC (5q21-22), and DCC (18q21.1), have been
20   identified in regions showing LOH.
         In breast cancer, loss of heterozygosity has
been reported on chromosomes 1p, 1q, 3p, 6q, 8q, 9q,
11p, 13q, 15q, 16q, 17p, 17q, 18p, and 22q (Devilee et
al., Oncogene 6: 1705-1711 (1991), Sato et al., Cancer
25   Res. 51: 5794-5799 (1991), Larsson et al., Genes,
Chromosomes and Cancer 2: 191-197 (1990)). LOH on
chromosome 16q is of particular interest for several
reasons. First, three different groups have reported a
high incidence of LOH on chromosome 16q in cells
30   derived from breast cancer tumors: 40% (Devilee et
al., supra.), 51% (Sato et al., supra.), and 53%
(Larsson et al., supra.). Second, LOH occurs on
chromosome 16q in other tumors, including
hepatocellular carcinoma, prostate cancer, and Wilms'
35   tumor (Tsuda et al., Proc. Natl. Acad. Sci. USA 87:
691-694 (1990), Nishida et al., Int. J.Cancer 51: 862-
868 (1992), Carter et al., Proc. Natl. Acad. Sci. USA

3

87: 8751-8755 (1990), Bergerheim *et al.*, *Genes,*
Chromosomes and Cancer 3: 215-220 (1991), Maw *et al.*,
Cancer Res. 52: 3094-3098 (1992)). Moreover, reduced
relative gene copy number has been detected on
5   chromosome 16q in breast cancer, prostate cancer and
ovarian cancer using comparative genomic hybridization
(CGH) (Kallioneimi *et al*, *Science* 258: 818-821 (1992)).

The particular genetic aberrations found in a
particular breast tumor can have a profound effect on
10 the prognosis and optimal treatment. The life
expectancy of breast cancer patients varies from a
relatively short time to as much as forty years
(Rutqvist *et al.*, *Cancer* 55: 658-665 (1985)), and is at
least partially dependent upon the particular genetic
15 changes that give rise to the tumor. For example, LOH
at the MET locus on chromosome 7q is associated with
aggressive disease in one subset of breast tumors
(Bieche *et al.*, *Lancet* 339: 139-143 (1992)).
Metastases from breast to the lymph nodes is increased
20 when the NM23 tumor suppressor gene is lost from
chromosome 17q (Hennessy *et al.*, *J. Int. Nat'l Cancer*
Inst. 83: 281-285 (1991). Therefore, diagnosis and
treatment of breast cancer depend upon the availability
of accurate and efficient methods for determining the
25 genetic aberrations that give rise to the particular
cancer.

Cytogenetics is the traditional method for
detecting amplified or deleted chromosomal regions.
The resolution of cytogenetic techniques is limited,
30 however, to regions larger than approximately 10 Mb
(approximately the width of a band in Giemsa-stained
chromosomes) because of the complex packing of DNA into
the chromosomes. In complex karyotypes with multiple
translocations and other genetic changes, traditional
35 cytogenetic analysis is of little utility because
karyotype information is lacking or cannot be
interpreted. Teyssier, J.R., *Cancer Genet. Cytogenet.*,

4

37: 103 (1989). Furthermore conventional cytogenetic
banding analysis is time consuming, labor intensive,
and frequently difficult or impossible.

5   More recent methods for assessing the amount
of a given DNA sequence in a chromosome are more .
sensitive than the traditional cytogenetic techniques.
These methods employ cloned DNA or RNA probes that are
hybridized to chromosomal DNA. This method is
effective even if the genome is heavily rearranged so
10   as to eliminate useful karyotype information. However,
this Southern blotting method gives only a rough
estimate of the copy number of a DNA sequence, and does
not provide any information about the localization of
that sequence within the chromosome.

15   Comparative genomic hybridization (CGH) is a
more recent approach to identify the presence and
localization of amplified or deleted sequences. *See,*
Kallioniemi *et al., supra.* CGH, like Southern
blotting, reveals amplifications and deletions
20   irrespective of genome rearrangement. CGH has
advantages over Southern blotting, providing a more
quantitative estimate of copy number, and also
providing the location of the amplified or deleted
sequence in the normal chromosome. Where a deletion or
25   amplification is limited to the loss or gain of one
copy of a sequence, the CGH resolution is usually about
5-10 Mb. Furthermore, the sensitivity of both CGH and
Southern blotting techniques to contamination of tumor
samples by normal cells makes estimating the copy
30   number of particular chromosomal sequences within the
tumor cell population very difficult.

New techniques which provide increased
sensitivity, more precise localization of a deletion
associated with LOH and neoplastic growth are
35   particularly desirable for the diagnosis of cancer.
The present invention provides these and other
benefits.

5

SUMMARY OF THE INVENTION

The present invention provides methods and
compositions for detecting deletions associated with
neoplastic growth. The methods comprise contacting a
5   chromosome sample from a patient with a composition
consisting essentially of one or more probes each of
which binds selectively to a target polynucleotide
sequence on human chromosome 16 distal to 16q21 under
conditions in which the probe forms a stable
10 hybridization complex with the sequence. The
hybridization complex is then detected using standard
techniques, typically by FISH.

The probe is usually labeled with digoxigenin
or biotin, which is visualized by detecting a
15 fluorescent label, such as FITC. In preferred
embodiments, the probe comprises a polynucleotide
sequence from C52, C196 and C21.

The hybridization complex is usually detected
in a metaphase cell in a sample from a breast tumor.
20 In addition, reference specific to chromosome 16
centromere may be used.

To confirm a loss of heterozygosity, the
methods may also include detecting a mutant allele of a
gene in the deleted region of chromosome 16. For
25 instance, E-cadherin expression can be determined.
Other methods for detecting a mutant allele include
PCR.

Also provided are nucleic acid probes which
bind selectively to a polynucleotide sequence on human
30 chromosome 16 distal to 16q21. The probes are
typically labeled with digoxigenin or biotin.
Preferred probes comprise polynucleotide sequences from
C52, C196 and C21.

The invention further provides kits
35 comprising a compartment which contains a nucleic acid
probe which binds substantially a polynucleotide on
human chromosome 16 distal to 16q21. The probes are

6

preferably labeled as described above. The kits may
also comprise compounds for the detection of the
probes, such as Texas red Avidin and biotin-labeled
goat anti-avidin antibodies.

5   Definitions

A "chromosome sample" as used herein refers
to a tissue or cell sample prepared for standard *in
situ* hybridization methods described below. The sample
is prepared such that individual chromosomes remain
10   substantially intact and typically comprises metaphase
spreads or interphase nuclei prepared according to
standard techniques. The tissue sample from which the
chromosome sample is prepared is typically taken from a
patient suspected of having the disease associated with
15   the deletion being detected.

As used herein a "probe" is defined as a
polynucleotide (either RNA or DNA) capable of binding
to a complementary target cellular genetic sequence
through one or more types of chemical bonds, usually
20   through hydrogen bond formation. It will be understood
by one of skill in the art that probes will typically
substantially bind target sequences lacking complete
complementarity with the probe sequence depending upon
the stringency of the hybridization conditions. The
25   probes are preferably directly labelled as with
isotopes or indirectly labelled such as with biotin to
which a streptavidin complex may later bind. By
assaying for the presence or absence of the probe, one
can detect the presence or absence of the target.
30   Nucleic acid probes can be prepared by a variety of
methods known to those of skill in the art.

A "composition consisting essentially of one
or more probes each of which binds selectively to a
35   target polynucleotide sequence" refers to a collection
of one or more probes which bind substantially to the
target sequence and nowhere else in the target

7

chromosome or genome and which allow the detection of the presence or absence of the target sequence. Such a composition may contain other nucleic acids which do not materially affect the detection of the target

5   sequence. Such additional nucleic acids include reference probes specific to a sequence in the centromere in the chromosome.

"Bind(s) substantially" refers to complementary hybridization between an oligonucleotide

10   and a target sequence and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target polynucleotide sequence.

"Hybridizing" refers the binding of two

15   single stranded nucleic acids via complementary base pairing.

"Nucleic acid" refers to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise

20   limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides.

One of skill will recognize that the precise sequence of the particular probes described herein can

25   be modified to a certain degree to produce probes that are "substantially identical" to the disclosed probes, but retain the ability to bind substantially to the target sequences. Such modifications are specifically covered by reference to the individual probes herein.

30   The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 90% sequence identity, and more preferably at least 95%, compared to a reference sequence using the methods described below using

35   standard parameters.

Two nucleic acid sequences are said to be "identical" if the sequence of nucleotides in the two

8

sequences is the same when aligned for maximum
correspondence as described below. The term
"complementary to" is used herein to mean that the
complementary sequence is identical to all or a portion
5   of a reference polynucleotide sequence.

Sequence comparisons between two (or more)
polynucleotides are typically performed by comparing
sequences of the two sequences over a "comparison
window" to identify and compare local regions of
10 sequence similarity. A "comparison window", as used
herein, refers to a segment of at least about 20
contiguous positions, usually about 50 to about 200,
more usually about 100 to about 150 in which a sequence
may be compared to a reference sequence of the same
15 number of contiguous positions after the two sequences
are optimally aligned.

Optimal alignment of sequences for comparison
may be conducted by the local homology algorithm of
Smith and Waterman *Adv. Appl. Math.* 2: 482 (1981), by
20 the homology alignment algorithm of Needleman and
Wunsch *J. Mol. Biol.* 48:443 (1970), by the search for
similarity method of Pearson and Lipman *Proc. Natl.
Acad. Sci. (U.S.A.)* 85: 2444 (1988), by computerized
implementations of these algorithms.

25   "Percentage of sequence identity" is
determined by comparing two optimally aligned sequences
over a comparison window, wherein the portion of the
polynucleotide sequence in the comparison window may
comprise additions or deletions (i.e., gaps) as
30 compared to the reference sequence (which does not
comprise additions or deletions) for optimal alignment
of the two sequences. The percentage is calculated by
determining the number of positions at which the
identical nucleic acid base or amino acid residue
35 occurs in both sequences to yield the number of matched
positions, dividing the number of matched positions by
the total number of positions in the window of

9

comparison and multiplying the result by 100 to yield
the percentage of sequence identity.

Another indication that nucleotide sequences
are substantially identical is if two molecules
5   hybridize to the same sequence under stringent
conditions. Stringent conditions are sequence
dependent and will be different in different
circumstances. Generally, stringent conditions are
selected to be about 5° C lower than the thermal
10   melting point (Tm) for the specific sequence at a
defined ionic strength and pH. The Tm is the
temperature (under defined ionic strength and pH) at
which 50% of the target sequence hybridizes to a
perfectly matched probe. Typically, stringent
15   conditions will be those in which the salt
concentration is at least about 0.02 molar at pH 7 and
the temperature is at least about 60°C.

                 BRIEF DESCRIPTION OF THE FIGURES
20   Figure 1. Shows the results of an experiment
examining intratumoral heterogeneity in tumor B208.
            Figure 2. Results of deletion mapping on
chromosome 16q in 23 primary breast cancer samples.
Closed boxes indicate deleted loci and open boxes
25   indicate retained loci. Deletions were detected by
fluorescence *in situ* hybridization.
            DESCRIPTION OF THE PREFERRED EMBODIMENTS
            This invention is based on the use of
libraries of genomic probes form chromosome 16 in *in*
30   *situ* hybridization to chromosome samples comprising
interphase nuclei or metaphase spreads to detect
chromosomal deletions. Generally, the method consists
of two steps: 1) the creation of a mapped library of
probes from chromosome 16, and 2) the *in situ*
35   hybridization of those probes to a sample and
subsequent detection of hybridization frequency to
determine relative copy number of a particular

10

chromosomal region (typically those distal to 16q21) in
the sample. The methods are preferably carried out
using fluorophores to detect the presence of the
hybridization complexes. These techniques are
5   generally referred to as fluorescent *in situ*
hybridization or FISH.

    The mapped libraries of probes consist of a
set of probes which when hybridized to a normal
chromosome are distributed relatively uniformly across
10 the region of interest. The region typically is one
chromosome or a part of one chromosome.

    FISH provides high resolution mapping of
cancer-associated amplifications and deletions. As the
physical mapping involved in the human genome project
15 proceeds, more chromosomal probes are becoming
available. In addition, FISH can be applied to
archived material (Schofield *et al.*, *American Journal
of Pathology* 141, 1265-1269 (1992)) or to the extremely
small amounts of material available from fine needle
20 aspirates or in touch preparations. Therefore, FISH
adds substantially to the battery of techniques that
can be used in locating genes involved in cancer
progression.

Production of a Probe Library

25   To prepare a mapped library, a selected
chromosome (*e.g.*, chromosome 16) is isolated by flow
cytometry using methods well known to those of skill in
the art. Briefly, chromosomes are isolated from cells
blocked in metaphase by the addition *e.g.*, colcemid and
30 stained with two DNA-binding fluorescent dyes. The
stained chromosomes are then passed through a cell
sorter and isolated using bivariate analysis of the
chromosomes by size and base pair composition (*see*,
e.g., Blennow *et al.*, *Hum. Genet.* 90:371-374 (1992).
35 The chromosome is then digested with restriction
enzymes appropriate to give DNA sequences of at least
about 20 kb and more preferably about 40 kb.

11

Techniques of partial sequence digestion are well known
in the art. See, for example Perbal, *A Practical Guide
to Molecular Cloning* 2nd Ed., Wiley N.Y. (1988). The
resulting sequences are ligated with a vector which
5   contains a resistance marker. The vector is
transfected into and propagated in the appropriate
host. Exemplary vectors suitable for this purpose
include cosmids, yeast artificial chromosomes (YACs),
bacterial artificial chromosomes (BACs) and P1 phage.
10 Typically, cosmid libraries are prepared. The cosmid
library then consists of single clones of the
transfected bacteria.

         While it is possible to generate cosmid
libraries, as described above, libraries spanning
15 entire chromosomes are available commercially
(Clonetech, South San Francisco, CA) or from the Los
Alamos National Laboratory.
<u>Mapping of Probe Library</u>

         Once a probe library is constructed, a subset
20 of the probes is physically mapped on the selected
chromosome, in this case chromosome 16. A number of
techniques can be used to obtain a mapped library.

         For instance overlapping sequences in a large
number of randomly selected cosmid clones can be
25 identified by unique restriction enzyme
"fingerprinting" and their assembly into overlapping
sets of clones. Such techniques have been used to map
various nonhuman genomes (see, e.g., Olson et al.,
Proc. Natl. Acad. Sci. USA 83:7826 (1986) and Coulson
30 et al., *Proc. Natl. Acad. Sci. USA* 83:7821 (1986).

         In addition, identification of overlapping
clones by hybridization with RNA probes can be used.
Cloning vectors are now available in which the
promoters for a single polymerase, or for two different
35 polymerases, lie adjacent to a cloning site.
Transcription with any of the available polymerases
enables one to produce large quantities of high-

12

specific activity RNA probes which correspond to either the coding or the non-coding strands (see, e.g. Wahl et al., Methods in Enzymology 152:572 (1987)). See also U.S. Patent No. 5,219,726.

5   Alternately, FISH and digital image analysis can be used to localize cosmids along the desired chromosome. This method is described, for instance, in Lichter et al., Science, 247:64-69 (1990). Briefly, the clones are mapped by FISH to metaphase spreads from
10   normal cells using e.g., FITC as the fluorophore. The chromosomes are counterstained by a stain which stains DNA irrespective of base composition (e.g., propidium iodide or DAPI), to define the outlining of the chromosome. The stained metaphases are imaged in a
15   fluorescence microscope with a polychromatic beam-splitter to avoid color-dependent image shifts. The different color images are acquired with a CCD camera and the digitized images are stored in a computer. A computer program is then used to calculate
20   the chromosome axis, project the two (for single copy sequences) FITC signals perpendicularly onto this axis, and calculate the average fractional length from a defined position, typically the p-telomere. When the p-telomere is used as the reference point the position
25   of each clone is expressed in units called FLpter.
          For use in the present invention, a mapped library of probes for chromosome 16 is described by Stallings et al., Genomics 13: 1031-1039 (1992).
Probe Preparation
30   The cosmid probes must be labeled for use in in situ hybridization. The probes may be detectably labeled prior to the hybridization reaction. Alternatively, a directly detectable label which binds to the hybridization product may be used. Such
35   detectable group can be any material having a detectable physical or chemical property and have been well-developed in the field of immunoassays. Thus a

13

label is any composition detectable by spectroscopic,
photochemical, biochemical, immunochemical, or chemical
means. Useful labels in the present invention include
32P, fluorescent dyes, electron-dense reagents, enzymes
5   (as commonly used in an ELISA). Examples of labels
which are not directly detected but are detected
through the use of directly detectable label include
biotin and dioxigenin as well as haptens and proteins
for which labeled antisera or monoclonal antibodies are
10   available.

The particular label used is not critical to
the present invention, so long as it does not interfere
with the *in situ* hybridization of the probe. In
addition the label must be detectible in as low copy
15   number as possible thereby maximizing the sensitivity
of the assay and yet be detectible above any background
signal. Finally, a label must be chosen that provides
a highly localized signal thereby providing a high
degree of spatial resolution when physically mapping
20   the probe against the chromosome. In a preferred
embodiment, the label is digoxigenin-11-dUTP or biotin-
14-dATP, which are then detected using fluorophores.

The labels may be coupled to the probes in a
variety of means known to those of skill in the art.
25   In a preferred embodiment the probe will be labeled
using nick translation or random primer extension
(Rigby, *et al. J. Mol. Biol.*, 113: 237 (1977) or
Sambrook, *et al.*, *Molecular Cloning - A Laboratory
Manual*, Cold Spring Harbor Laboratory, Cold Spring
30   Harbor, N.Y. (1985)).

In situ Hybridization

Standard *in situ* hybridization techniques are
used to probe a given sample. Several guides to the
techniques are available, *e.g.*, Gall *et al. Meth.*
35   *Enzymol.*, 21:470-480 (1981) and Angerer *et al.* in
Genetic Engineering: Principles and Methods Setlow and
Hollaender, Eds. Vol 7, pgs 43-65 (plenum Press, New

14

York 1985). Briefly, cells, either as single cell
suspensions or as tissue preparation are deposited on
solid supports such as glass slides and fixed by
choosing a fixative which provides the best spatial
5   resolution of the cells and the optimal hybridization
efficiency.

Generally, *in situ* hybridization comprises
the following major steps: (1) fixation of tissue or
biological structure to analyzed; (2) prehybridization
10   treatment of the biological structure to increase
accessibility of target DNA, and to reduce nonspecific
binding; (3) hybridization of the mixture of nucleic
acids to the nucleic acid in the biological structure
or tissue; (4) posthybridization washes to remove
15   nucleic acid fragments not bound in the hybridization
and (5) detection of the hybridized nucleic acid
fragments. The reagent used in each of these steps and
their conditions for use vary depending on the
particular application.

20   In some applications it is necessary to block
the hybridization capacity of repetitive sequences. In
this case, human genomic DNA is used as an agent to
block such hybridization. The preferred size range is
from about 200 bp to about 1000 bases, more preferably
25   between about 400 to about 800 bp for double stranded,
nick translated nucleic acids.

Hybridization protocols for the particular
applications disclosed here are described in detail
below. Suitable protocols are described in Pinkel *et*
30   *al. Proc. Natl. Acad. Sci. USA*, 85:9138-9142 (1988) and
in EPO Pub. No. 430,402.

Typically, it is desirable to use dual color
FISH, in which two probes are utilized, each labelled
by a different fluorescent dye. A test probe that
35   hybridizes to the region of interest is labelled with
one dye, and a control probe that hybridizes to a
different region is labelled with a second dye. A

nucleic acid that hybridizes to a stable portion of the chromosome of interest, such as the centromere region, is often most useful as the control probe. In this way, differences between efficiency of hybridization

5 from sample to sample can be accounted for.

As noted above, the methods of the invention are particularly useful in detecting LOH. In this case, LOH regions distal to 16q21are associated with neoplastic growth. As shown below, detection of the

10 presence of a mutant allele can be used to confirm that the deletion leads to LOH. One of skill will recognize a number of methods for detecting a mutant allele. Examples include southern hybridization, PCR, restriction length polymorphisms, or detection of the

15 gene product using, e.g., immunological techniques.

Sample Preparation

The FISH methods for detecting chromosomal amplifications and deletions associated with cancer described herein can be performed on nanogram

20 quantities of the subject nucleic acids. Paraffin embedded tumor sections can be used, as can fresh or frozen material. Because FISH can be applied to the limited material, a number of types of samples can be used. For instance, small biopsy tissue samples from

25 tumors can be used for touch preparations (see, e.g., Kallioniemi, A. et al., Cytogenet. Cell Genet. 60: 190-193 (1992)). Small numbers of cells obtained from aspiration biopsy or cells in bodily fluids (e.g., blood, urine, sputum and the like) can also be

30 analyzed.

EXAMPLE 1

CGH studies suggested that an LOH on chromosome 16q is associated with breast cancer, and that this LOH is due to a physical deletion. To map

35 the location of this LOH in primary breast tumors, we performed dual color FISH (Matsumara et al., Cancer Res. 52: 3474-3477 (1992)) with probes distributed

16

along chromosome 16q. A control probe that hybridizes
to the centromere region of chromosome 16 was labelled
with one color dye, while the test probes were labelled
with a second color dye. Cells were scored to
5   determine if the test probe produced fewer signals than
the centromere probe in greater than 50% of all cells.
Reduced copy number of the test probe target regions
relative to the centromere region is strongly
correlated with loss of heterozygosity at other loci.
10   FISH was performed with probes obtained from
Los Alamos. Precise mapping and contiguity of the
probes was then confirmed prior to their use in
analyzing samples.
Methods
15   Cosmids or yeast artificial chromosomes
(YACs) that contain sequences derived from regions
distributed along chromosome 16q were labeled with
either digoxigenin-11-dUTP or biotin-14-dATP by
nick-translation. PCR based screening of a YAC library
20   was performed using primer pairs for exons 5 and 6 of
the human E-cadherin gene to produce an E-cadherin
specific probe. Biotinylation was carried out using a
nick translation kit (Gibco BRL); the reaction was
carried out for 90 min at 15°C. For digoxigenin
25   labeling, digoxigenin-11-dUTP (Boehringer Mannheim) was
used instead of biotin-14-dATP.
The sample slides were denatured by
incubating in denaturing solution (70% formamide, 2X
SSC, pH 7.0) at 73°C for 2.5 min. After denaturation,
30   slides were digested with proteinase K (Sigma) at 37°C
for 75 min (0.1 μg/ml for lymphocytes, 0.5 μg/ml for
cultured cells, and 1.0 μg/ml for clinical samples, in
20mM Tris, 2mM $CaCl_2$, pH7.5).
Lymphocytes were used as normal control.
35   Cell lines were used for deletion mapping of the E-
cadherin gene. Slides were prepared as follows: Touch
preparations were prepared from fresh primary tumors by

17

pressing the cut surface of each tumor to a several
clean microscope slides. Cells adhering to the slides
were air dried and fixed for 5 min each in three
changes of freshly prepared Carnoy's solution (3:1 =
5   methanol: acetic acid). Cell suspensions were prepared
from fresh tumors by mechanically disaggregating the
tumors. The suspended cells were incubated in
hypotonic KCl (75mM) for 15 min at 37°C, fixed in
Carnoy's solution, and dropped onto slides. Cell lines
10   were cultured in appropriate culture media in
humidified chamber under 5% CO_2 at 37°C. Metaphase
spreads from breast cancer cell lines were made after
treating these cells with colcemid (0.25µg/ml) for 3
hours. After hypotonic KCl treatment for 15 min at
15   37°C, the cells were fixed in Carnoy's solution,
dropped onto slides and air dried. All slides were
stored at -70°C until use under nitrogen gas.

        Following denaturation of the samples, a
probe mixture was applied to a warmed slide according
20   to standard procedures. The probe mixture contained 20
ng of labelled probe, 0.2 ng of chromosome 16-specific
a-satellite probe (control probe), and, as blocking
DNA, either 10 µg of human placental DNA (Sigma) for
cosmid probes or 2µg of human Cot 1 DNA (Gibco BRL) for
25   a YAC probe. The probes and carrier DNA were in 50%
formamide, 2X SSC, and 10% dextran sulfate at pH7.0).

        The probes were hybridized to the samples for
18-36 hours under a cover slip at 37°C in a moist
chamber. After incubation, the slides were washed 3
30   times in washing buffer (50%) formamide, 2 X SSC,
pH7.0), twice in 2X SSC, once in 0.1X SSC, each wash
being carried out for 10 min at 45°C. A final wash in
4X SSC was carried out for 5 min at room temperature.

        A three-step sandwich procedure was used to
35   immunocytochemically stain the slides after
hybridization and washing. Hybridized biotin-labeled
probes were detected with Texas Red avidin (Vector

18

Laboratories), biotin-labeled goat anti-avidin antibody
(Vector Laboratories), and finally Texas Red avidin.
The digoxigenin-labeled probes were detected with FITC-
labeled sheep anti-digoxigenin antibody (Boehringer

5   Mannheim) and finally FITC-labeled rabbit anti-sheep
antibody (Vector Laboratories).

The staining protocol was carried out at room
temperature. Before staining, the slides were
preblocked in 1% bovine serum albumin (BSA) in 4X SSC

10 for 5 min. The first staining step consisted of 2μg/ml
of Texas Red avidin in 1% BSA/4X SSC. The slides were
then washed sequentially in 4X SSC, 4X SSC/0.1% Triton
X-100, 4X SSC, and finally in PN (a mixture of 0.1M
NaH$_2$PO$_4$ and 0.1M NaH$_2$PO$_4$ (pH 8.0), and 0.1% Nonidet P-40)

15 for 10 min each. The slides were then preblocked by
incubating in PNM (5% Carnation dry milk/0.02% sodium
azide in PN) for 5 min.

The second step consisted of incubating the
slides in PN containing 20 μg/ml FITC-labeled sheep

20 anti-digoxigenin antibody and 5 μg/ml biotin-labeled
anti-avidin antibody. Three 10 minute washes in PN
followed. The slides were then preblocked by
incubating in PMN for 5 min.

The third step was done with FITC-labeled

25 rabbit anti-sheep antibody (50:1 dilution) and 2 μg/ml
Texas Red avidin in PN. After three PN washes, nuclei
were counterstained with 0.2 μM 4′, 6-diamidino-2-
phenylindole (DAPI) in an antifade solution.

Results

30   Cells were scored to determine if the test
probe produced less signal relative to that from the
control centromere region probe in greater than 50% of
all cells. Figure 1 shows the results of this relative
copy number analysis at 5 sites along chromosome 16q in

35 a primary breast cancer, B208. Cosmid probes C89 and
C25, the two probes most proximal to the centromere,
gave a signal nearly equal to the control signal in 95%

19

and 89%, respectively, of the cells examined. In
contrast, the test probes C52, C196, and C21 (which lie
more distal than the C89 and C25 probes) gave a much
lower signal relative to the control probe in nearly
5   all of the cells.

Twenty-three independent primary breast
cancers were analyzed in this way and the results are
shown in Figure 2. Deletions of all material distal to
16q13 were detected in 10 tumors and deletions of all
10   material distal to 16q21 were detected in an additional
three tumors. Thus, all tumors showing physical
deletions on 16q appeared to be deleted distal to
16q21. Eight of the 13 tumors scored as deleted on 16q
appeared to carry only one copy of chromosome 16q and 5
15   appeared to have more than one. Thus, all tumors
retained at least one apparently intact copy of
chromosome 16q.

These results strongly suggest that the
deletions on 16q are unmasking recessive mutations in a
20   putative tumor suppressor gene encoded on 16q. Since
all the deletions were mapped distal to 16q21, we
reasoned that the suppressor gene must be located just
distal to 16q21.

25   EXAMPLE 2

The E-cadherin gene maps to 16q22.1, just
distal to the common deletion at 16q21 shown above. E-
cadherin (also known as Arc-1, uvomorulin, L-CAM, and
cell-CAM-120/80) is a 120 kD cell surface glycoprotein
30   adhesion molecule that connects epithelial cells in a
calcium-dependent homophilic fashion (Nagafuchi et al.,
Nature 329: 341-343 (1987); Ozawa et al., Cell 63:
1033-1038 (1990)). Several recent findings have
suggested that E-cadherin may be a tumor suppressor
35   gene. For example, loss of E-cadherin expression is
associated with dedifferentiation and invasiveness of
cancer cells (Umbas, et al., Cancer Research 52: 5104-

20

5109 (1992); Giroldi et al., *Cancer and Metastasis
Reviews* 12: 29-37 (1993); Dorudi et al., *Am J Pathol*
142: 981-986 (1993); Mayer, et al., *Cancer Res* 53:
1690-1695 (1993); Bringuier et al., *Cancer Research*
5   53: 3241-3245 (1993); Gamallo et al., *Am J Pathol* 142:
987-993 (1993); Oka et al., *Cancer Research* 53: 1696-
1701 (1993)). E-cadherin expression is inversely
correlated with the grade of *in vitro* invasiveness in
bladder, pancreas, lung, and breast cancer cells lines
10 (Frixen et al., *J. Cell Biology* 113: 173-185 (1991)).
Transfection of human breast cancer cell lines with
mouse E-cadherin cDNA reduced invasiveness of the cell
lines *in vitro* (Frixen et al., *supra.*). A single
amino acid substitution in one of the Ca^{2+} binding sites
15 of E- cadherin causes loss of cell-cell adhesion (Ozawa
et al, *Cell* 63: 1033-1038 (1990)). Also, E-cadherin
has regions of homology with *fat*, a tumor suppressor
gene in *Drosophila* (Mahoney et al., *Cell* 67: 853-868
(1991)).

20   We tested the hypothesis that the 16q
deletions are indeed unmasking mutations in the E-
cadherin gene by detailed analysis of E-cadherin copy
number, sequence, and expression in 11 breast cancer
cell lines. FISH analysis of relative copy number at
25 E-cadherin in metaphase spreads and interphase nuclei
was accomplished using a YAC selected from a genomic
library.

            Ten established breast cancer cell lines
(BT468, BT474, CAMA1, HS578T, MCF-7, MDA134, MDA361,
30 MDA435, SKBR3, and T47D) were obtained from the
American Type Culture Collection (ATCC). One cell
line, 600MPE, was provided by Dr. Helene Smith.
Methods
            The test probe for FISH analysis of the E-
35 cadherin locus was a YAC clone that contains the human
E-cadherin gene. The appropriate YAC (clone 281D8)
obtained by polymerase chain reaction (PCR) screening

21

of the CEPH YAC library. Two primer sets were used to
amplify 145 bp within exon 5 of E-cadherin and 121 bp
within exon 6 of the gene.

       Primers used to amplify various exons within

5   the E-cadherin gene were as follows:

EX1-2.A: 5'-CGGAAGTCAGTTCAGACTCC-3',

EX1-2.B: 5'-CCTCTCTCCAGCTCC-3',

EX3.A: 5'-TTTGAAGATTGCACC-3',

EX3.B: 5'-CCACTGTATTCAGCG-3',

10 EX4-5.A: 5'-CGGACGATGATGTGAACAC C-3',

EX4-5.B: 5'-TGAAGATCGGAGGATTATCG-3',

EX5.A: 5'-CCAACAAAGACAAAGAAGGC-3',

EX5.B: 5'-TATGTGGCAATGCGTTCTCTATCC-3',

EX6.A: 5'-TCTCTCACGCTGTGTCATCCAACG-3',

15 EX6.B: 5'-ACAGACCCCTTAAAGACCTCCT GG-3',

EX7-8.A: 5'-CGGACGATGATGTGAACACC-3',

EX7-8.B: 5'-TGAAGATCGGAGGATTATCG-3',

EX9.A: 5'-AGAACGAGGCTAACG-3',

EX9.B: 5'-TCACTGGATTTGTGG-3',

20 EX10.A: 5'-TTCTACACGTAGCAGTGACG-3',

EX10.B: 5'-GTTCCATAAATGTGTCTGGC-3',

EX11.A: 5'-CGGATTTGGAGAGACACTGC-3',

EX11.B: 5'-TATGATTAGGGCTGTGTACG-3',

EX12.A: 5'-CAGTTGCTACTGGAACAGGG-3',

25 EX12.B: 5'-TGGGTCGTTGTACTGAATGG-3',

EX13.A: 5'-AGCTCATGGATAACC-3',

EX13.B: 5'-TCCTCCAAGAATCCC-3',

EX14.A: 5'-CTGCTCTTGCTGTTTCTTCG-3',

EX14.B: 5'-TGGTCCTCTTCTCCGCCTCC-3',

30 EX14-15.A: 5'-CTGCTCTTGCTGTTTCTTCG-3',

EX14-15.B: 5'-TCCAATTTCATCGGG-3',

EX16.A: 5'-GAAAGCGGCTGATACTGACC-3',

EX16.B: 5'-ACGTGATTTCTGCATTTCCC-3'.

Pairs of primers (EX5.A, EX5.B and EX6.A, EX6.B) were

35 used for screening of the YAC library.

       E-cadherin expression was assessed by
indirect immunofluorescence staining using an antibody

22

against the extracellular portion of E-cadherin.
Cultured breast cancer cells were scraped from the wall
of plastic culture flasks, washed 3 times in phosphate
buffered saline and adjusted to 1x10^6 cells in 100 μl of
5   PBS. Ten microliters of mouse anti-E-cadherin
monoclonal antibody E9 (*J. Cell. Biochem.* 34:187
(1987)) was added to each cell suspension and incubated
for 30 min at 37°C. The cells were then washed 3 times
in PBS, incubated in FITC-labeled sheep anti-mouse Ig
10   antibody (Boehringer Mannheim) for 30 min at 37°C,
     washed 3 times in PBS and resuspended again in 50 μl of
PBS. Fluorescence of the cells were observed under a
Zeiss fluorescence microscope.
          The structural integrity and expression of
15   the non-deleted E-cadherin allele in the breast cancer
cell lines was assessed by non-isotopic SSCP analysis
(Yap *et al.*, *Nucleic Acids Res* 20: 145 (1992);
Ainsworth *et al.*, *Nucleic Acids Res* 19: 405-406 (1991);
Ballhausen *et al.*, *Applied and Theoretical*
20   *Electrophoresis* 3: 129-131 (1993)) and reverse
transcriptase mediated PCR (RT-PCR).
          The full length sequence of both the human
and murine E-cadherin cDNA have recently been reported
(Bussemakers *et al.*, *Molecular Biology Reports* 17: 123-
25   128 (1993)). The genomic organization of the mouse
gene, including the intron-exon boundaries, has been
deduced (Ringwald *et al.*, *Nucleic Acids Research* 19:
6533-6539 (1991)). The intron-exon boundaries of the
human E-cadherin gene were deduced by comparison with
30   the mouse homolog, and PCR primers were designed to
amplify as much as possible of each of the predicted 16
exons. The PCR primers, designed using the program
Primer Detective (Lowe *et al.*, *Nucleic Acids Research*
18, 1757-1761 (1990)), allowed analysis of
35   approximately 80% of the E-cadherin coding sequence.
          The SSCP analysis was carried out as follows.
Ten μl of PCR product mixed with 2 μl of loading dye

23

(0.5% bromophenol blue and 0.5% xylene cyanol in formamide) was heat-denatured at 95°C for 5 min and cooled on ice. The samples were then electrophoresed on 0.8 mm thick non-denaturing 6% or 10%

5   polyacrylamide gels (acrylamide: bisacrylamide=49:1 or 19:1) with or without 2.5% or 5% glycerol in 0.5 x TBE buffer at 4-5 V/cm for 3 to 8 hours at room temperature. Gels were stained for 15 min in 0.5X TBE containing 5 µg/ml ethidium bromide. Double stranded

10 and single stranded DNA were visualized on UV transilluminator. Genomic DNA extracted from normal human lymphocytes was used as control template DNA.

        Reverse transcriptase-mediated PCR was ·
carried out as follows. Total cellular RNA was

15 extracted from approximately 5x10^6 cells using Tri Reagent (Molecular Research Center, Inc.), following the manufacturer's protocol. The structural integrity of the isolated RNA was confirmed by electrophoresis of the RNA through a 1.2% formaldehyde gel. Five

20 micrograms of total RNA from cell lines MCF7, 600MPE, CAMA1 and 578T was converted to first strand cDNA using the Preamp cDNA synthesis kit (BRL). RT-PCR was performed on 1 µl of the synthesized cDNA template in a 10 ml reaction volume containing 1 mM Mg^{++}, 50 mM Tris-

25 Cl, pH 8.3, 0.25 mg/ml BSA, 0.5% Ficoll 400, 1 mM tartrazine dye; 10 pM each of primers 5U and 10L and 0.5U Taq polymerase (Perkin-Elmer). The RT-PCR reactions were performed in a Idaho Technology Air Thermo-Cycler 1605 using 10 ml capillary tubes (Idaho

30 Technology) under the following conditions: 1 minute hard denaturation at 94°C followed by 30 cycles of (0 sec. @ 94°C, 0 sec. @ 58°C, 25 sec. @ 70°C and a 2 minute final extension at 70°C. The PCR products were loaded directly into a 1.2% agarose gel for size

35 fractionation.

Results

24

All cell lines tested showed at least one E-cadherin locus, as determined by FISH analysis. The copy number of E-cadherin was reduced relative to the control centromere region copy number in four of the

5   eleven cell lines (MDA134, MDA435, SKBR3, and 600MPE). Another line, CAMA1, had the same number of E-cadherin and centromere hybridization signals. This cell line has two 16q isochromosomes plus two additional chromosomes that have chromosome 16 centromeres and

10   lack E-cadherin signals. This cell line was scored as deleted at the E-cadherin locus, using the rationale that one E-cadherin allele was lost from this cell line through deletion, and the other E-cadherin allele was duplicated through formation of an isochromosome 16q,

15   and duplicated again as the cell became tetraploid for chromosome 16q. Significantly, the three primary tumors showing reduced relative copy number distal to 16q21 (Fig. 2) also showed reduced relative copy number at the E-cadherin locus.

20   Indirect immunofluorescence staining using an antibody directed against the extracellular portion of E-cadherin revealed that all five cell lines having reduced relative copy number at the E-cadherin locus failed to express E-cadherin. Five of the six cell

25   lines in which the relative copy number of E-cadherin was not significantly reduced expressed E-cadherin. These results are summarized in Table 1, and confirm previous reports (Frixen, et al., supra.; Sommers, et al., Cell Growth and Differentiation 2: 365-372

30   (1991)).

The failure of the five cell lines having reduced relative E-cadherin gene copy number to express detectable E-cadherin protein on the cell surface suggested that the remaining E-cadherin allele

35   contained inactivating mutations. The structural integrity and expression of the remaining E-cadherin allele were analyzed by non-isotopic SSCP analysis and

25

reverse transcriptase mediated polymerase chain
reaction (RT-PCR).

All five cell lines that have reduced
relative copy number at the E-cadherin locus were found
5   to carry DNA mutations in the remaining E-cadherin
allele, or did not produce detectable E-cadherin mRNA.
MDA134 and SKBR3 were found to have deletions in the E-
cadherin gene, as PCR failed to amplify sequences in
these exons. Specifically, exon 6 is deleted in MDA134
10   and exons 3 through 12 are deleted in SKBR3. SSCP
analysis of 600MPE, revealed no mutation or deletions.
However, RT-PCR of mRNA from this line, using primers
designed to amplify from exon 5 to exon 10, revealed a
deletion of approximately 180 bp. The deleted fragment
15   was subcloned and sequenced, revealing a deletion of
the 183 bp exon 9.

This result suggested the presence of an
intronic mutation that interferes with the normal
processing of the E-cadherin message. To characterize
20   the mutation, genomic DNA from normal cells and from
600MPE were amplified using primers located in exons 8
and 9 (see above) and the resulting approximately 1 kb
fragments were cloned and sequenced. This analysis
revealed a 22 bp deletion in intron 8 of 600MPE that
25   eliminates the splice site acceptor sequence.

The cell lines CAMA1 and 57ST were found to
be normal in all exons by SSCP. However, RT-PCR
analysis showed that these lines did not express
detectable E-cadherin mRNA. This suggests that these
30   lines carry mutations that affect either the expression
or stability of the E-cadherin message.

Exonic DNA in all of the cell lines that
express E- cadherin appeared normal by SSCP analysis,
except for one mobility shift detected in DNA fragments
35   amplified from exon 14 in BT468. DNA sequencing
analysis of a PCR amplification product from this exon
revealed a single base substitution (C→T) at base

position 2526. However, this substitution did not change the amino acid Asn[751] encoded by this codon (AAC→AAT), so this mutation had no physiological effect. A single polymorphism was detected in exon 13
5   at base position 2170 (C or T). This substitution does not change the coding amino acid Ala[692] (GCC→GCT). Two of five E-cadherin positive cell lines (BT474 and MDA361) and 7 of 10 normal cell lines were heterozygous for this polymorphism. However, all six E-cadherin
10   negative cell lines had only one allele at the site of the polymorphism, as expected for cells that had lost one E-cadherin allele through physical deletion.

In summary, these data suggest that E-cadherin is a tumor suppressor gene that is inactivated
15   in breast cancer through deletion of one allele leading to LOH on 16q, and mutation of the other allele. The frequency at which E-cadherin is inactivated in primary breast cancer has not yet been assessed at the genetic level. However, estimates of LOH in mixed breast
20   cancers range from 40 to 50% (Devilee et al., *Oncogene* 6: 1705-1711 (1991); Sato et al., *Cancer Res.* 51: 5794-5799 (1991); Larsson et al., *Genes, Chromosomes and Cancer* 2: 191-197 (1990)). In addition, immunohistochemical analyses of E-cadherin expression
25   showed that expression is reduced in 71% of T3/T4 tumors, 74% of tumors with extensive lymph node metastasis, and in 86% of tumors with distant metastases (Oka et al., *Cancer Res.* 53: 1696-1701 (1993)). Loss of E-cadherin has been reported in
30   advanced cancers of the prostate (Umbas et al., *Cancer Res.* 52: 5104-5109 (1992); Giroldi et al., *Cancer and Metastasis Reviews* 12: 29-37 (1993)), colon (Dorudi et al., *Am J Pathol* 142: 981-986 (1993)), stomach (Mayer et al., *Cancer Res* 53: 1690-1695 (1993)) and bladder
35   (Bringuier et al., *Cancer Res.* 53: 3241-3245 (1993)) and abnormal E-cadherin expression in bladder cancer correlates with shorter survival (Bringuier et al.,

supra.). Thus, loss of E-cadherin function leading to reduced cell adhesion seems likely to play an important role in the development of invasive potential in a high frequency of breast cancers and other solid tumors.

5   The frequency of tumors progressing through loss of E-cadherin function may be even higher than suggested by analyses focused solely on E-cadherin, because dysfunction can also be caused by disruption of the E-cadherin-catenin complex that couples E-cadherin 10 to the cellular cytoskeleton (Ozawa *et al.*, *Proc. Natl. Acad. Sci. USA* 87: 4246-4250 (1990)). This may be caused by loss of function in α-, β- or γ-catenin (Shimoyama *et al.*, *Cancer Res.* 52: 5770-5774 (1992)), or by phosphorylation of the E-cadherin-catenin complex 15 by activation of the v-src oncogene (Behrens *et al.*, *J. Cell Biol.* 120: 757-766 (1993)). Of course, other genes on 16q also may play a role in cancer progression since the LOH and FISH studies described here and elsewhere frequently involve regions distal to E- 20 cadherin. The Cell Matrix Adhesion Regulator (CMAR) gene at 16q24 (Koyama *et al.*, *Genomics* 16: 264-265 (1993)) is a candidate second tumor suppressor gene on chromosome 16q. CMAR acts to increase cell adhesion to components of the extracellular matrix (Pullman *et al.*, 25 *Nature* 356, 529-532 (1992)).

      All of the references cited above are hereby incorporated by reference. For the purposes of clarity and understanding, the invention has been described in these examples and the above disclosure in some detail. 30 It will be apparent, however, that certain changes and modifications may be practiced within the scope of the appended claims.

28.

Table 1   Expression, gene copy number, and aberrations in the human E-cadherin gene in breast cancer cell lines.

Cell lines	E-cadherin Protein Expression	Predominant population 16 centromere/E-cadherin YAC copies/cell	E-cadherin mRNA Expression	Aberrations in genomic DNA of the E-cadherin gene detected by SSCP and RT-PCR
BT468	+	2/4	+	ND
BT474	+	5/5	+	ND
MCF-7	+	5/6	+	ND
MDA361	+	4/4	+	ND
T47D	+	3/3	+	ND
CAMA1	-	4/4*	+	ND
HS578T	-	2/2	-	ND
MDA134	-	2/1**	-	Deletion (exon 6)
MDA435	-	4/2**	-	ND
SKBR3	-	4/2**	-	Deletion (exon 3-exon 12)
600MPE	-	2/1**	+	Splice acceptor site deletion (intron8)

*Isochromosomes   **Deletionon of the E-cadherin gene   ND: Not detected

29

WHAT IS CLAIMED IS:

1.   A method of detecting a deletion
associated with neoplastic growth, the method
comprising:

5   contacting a chromosome sample from a patient
with a composition consisting essentially of one or
more labeled nucleic acid probes each of which binds
selectively to a polynucleotide sequence on human
chromosome 16 distal to 16q21 under conditions in which
10   the probe forms a stable hybridization complex with the
sequence; and
            detecting the hybridization complex.

            2.   The method of claim 1, wherein the label
15   is selected from the group consisting of digoxigenin
and biotin.

            3.   The method of claim 1, wherein the step
of detecting the hybridization complex is carried out
20   by detecting a fluorescent label.

            4.   The method of claim 3, wherein the
fluorescent label is FITC.

25   5.   The method of claim 1, wherein the probe
is selected from the group consisting of polynucleotide
sequences from C52, C196 and C21.

            6.   The method of claim 1, wherein the
30   hybridization complex is detected in a metaphase cell.

            7.   The method of claim 1, wherein the
sample is from a breast tumor.

35   8.   The method of claim 1, further
comprising detecting E-cadherin expression in a sample
from the patient.

9.   The method of claim 1, further
comprising contacting the sample with a reference probe
which binds selectively to a sequence in the centromere
of chromosome 16.

5

10.   A method of detecting a deletion
associated with neoplastic growth, the method
comprising,

contacting a chromosome sample with a
10 composition consisting essentially of one or more
labeled nucleic acid probes each of which binds
selectively to a polynucleotide sequence on human
chromosome 16 distal to 16q21 under conditions in which
the probe forms a stable hybridization complex with the
15 sequence;

detecting the hybridization complex; and
detecting a mutant allele of a gene located
on human chromosome 16 distal to 16q21.

20   11.   The method of claim 10, wherein the
mutant allele is detected using PCR.

12.   The method of claim 10, wherein the
mutant allele is detected by detecting expression of
25 the gene.

13.   The method of claim 10, wherein the gene
is E-cadherin.

30   14.   A composition comprising a nucleic acid
probe which binds selectively to a polynucleotide
sequence on human chromosome 16 distal to 16q21.

15.   The composition of claim 14, wherein the
35 probe is labeled.

31

16. The composition of claim 14, wherein the label is selected from the group consisting of digoxigenin and biotin.

17. The composition of claim 14, wherein the probe is selected from the group consisting of polynucleotide sequences from C52, C196 and C21.

18. A kit for the detection of deletions associated with neoplastic growth, the kit comprising a compartment which contains a nucleic acid probe which binds selectively to a polynucleotide on human chromosome 16 distal to 16q21.

19. The kit of claim 18, wherein the probe is labeled.

20. The kit of claim 19, wherein label is selected from the group consisting of digoxigenin and biotin.

21. The kit of claim 18, wherein the probe is selected from the group consisting of polynucleotide sequences from C52, C196 and C21.
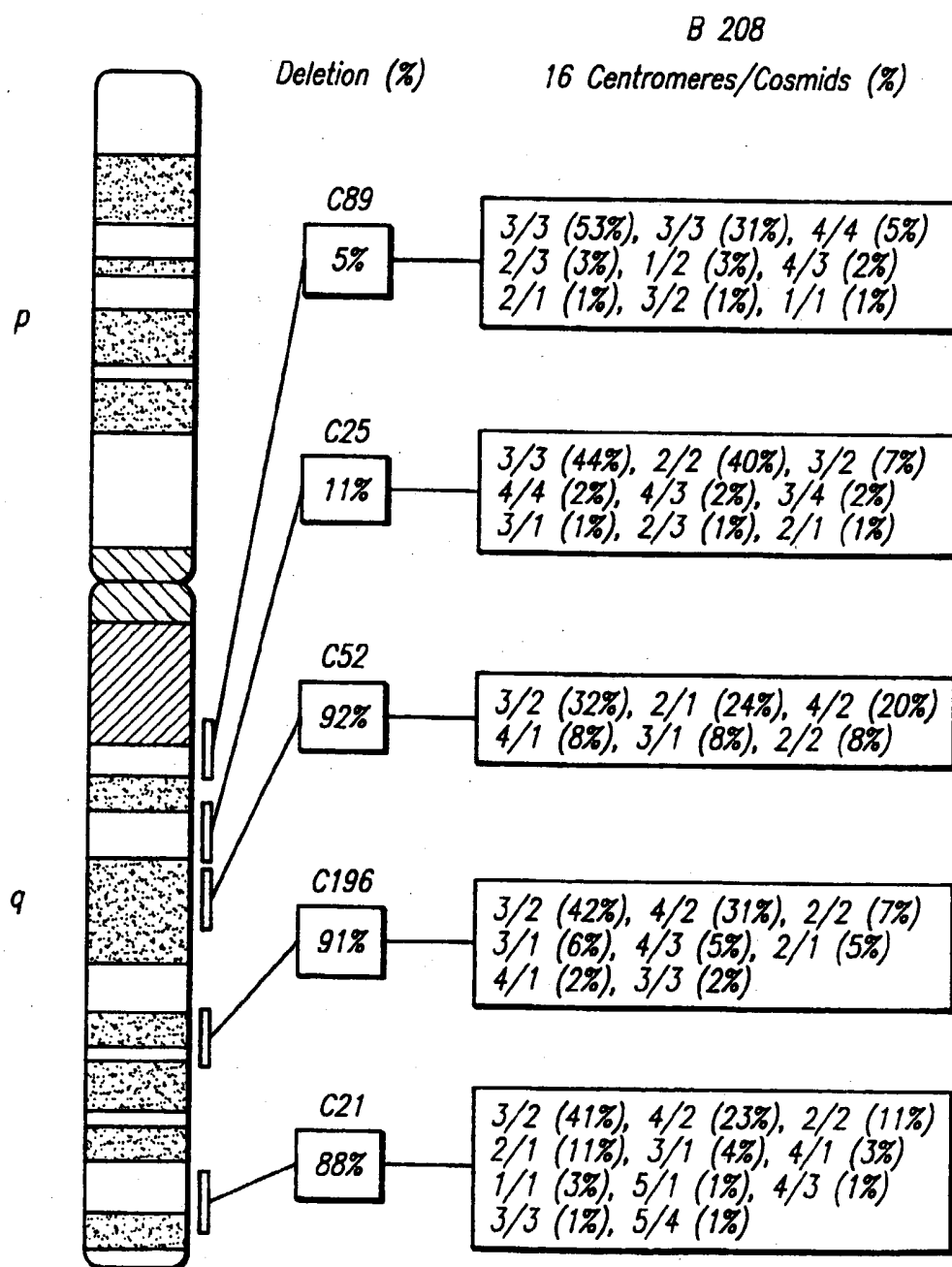
22. The kit of claim 18, further comprising a reference probe specific to chromosome 16 centromere.

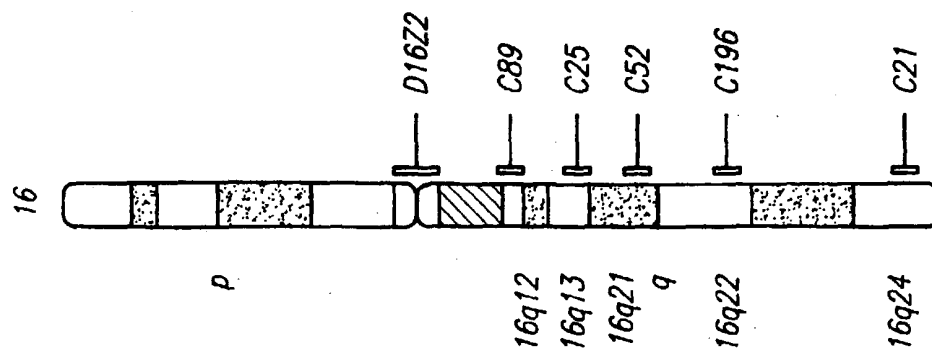23. The kit of claim 18, further comprising Texas red Avidin and biotin-labeled goat anti-avidin antibodies.
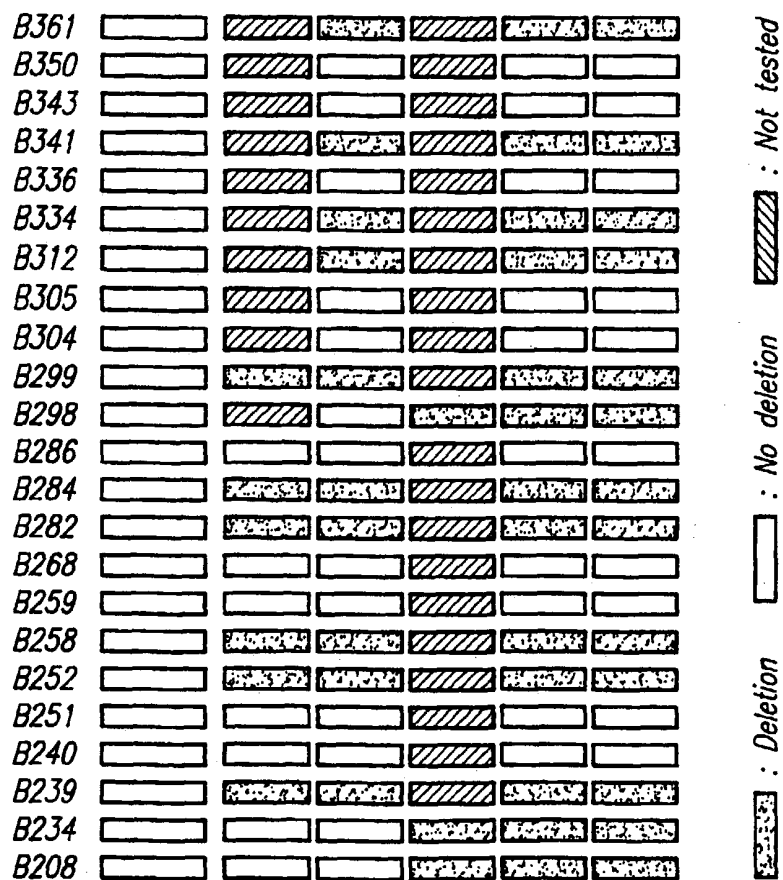
FIG. 1

| | Deletion (%) | B 208
16 Centromeres/Cosmids (%) |

C89 — 5%
3/3 (53%), 3/3 (31%), 4/4 (5%)
2/3 (3%), 1/2 (3%), 4/3 (2%)
2/1 (1%), 3/2 (1%), 1/1 (1%)

C25 — 11%
3/3 (44%), 2/2 (40%), 3/2 (7%)
4/4 (2%), 4/3 (2%), 3/4 (2%)
3/1 (1%), 2/3 (1%), 2/1 (1%)

C52 — 92%
3/2 (32%), 2/1 (24%), 4/2 (20%)
4/1 (8%), 3/1 (8%), 2/2 (8%)

C196 — 91%
3/2 (42%), 4/2 (31%), 2/2 (7%)
3/1 (6%), 4/3 (5%), 2/1 (5%)
4/1 (2%), 3/3 (2%)

C21 — 88%
3/2 (41%), 4/2 (23%), 2/2 (11%)
2/1 (11%), 3/1 (4%), 4/1 (3%)
1/1 (3%), 5/1 (1%), 4/3 (1%)
3/3 (1%), 5/4 (1%)

16

FIG. 2

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/68; C07H 21/02, 21/04
US CL : 435/6; 536/22.1,23.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6; 536/22.1,23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

BIOSIS, APS,CAS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GENOMICS, VOLUME 13, ISSUED 1992, CALLEN ET AL, "HIGH-RESOLUTION CYTOGENETIC-BASED PHYSICAL MAP OF HUMAN CHROMOSOME 16", SEE PAGES 1178-1185, SEE ENTIRE DOCUMENT.	1-13
Y	CANCER RESEARCH, VOLUME 50, ISSUED 15 NOVEMBER 1990, SATO ET AL, "ALLELOTYPE OF BREAST CANCER: CUMULATIVE ALLELE LOSSES PROMOTE TUMOR PROGRESSION IN PRIMARY BREAST CANCER", PAGES 7184-7189, SEE ENTIRE DOCUMENT.	1-13
Y	CANCER RESEARCH, VOLUME 53, ISSUED 01 APRIL 1993, OKA ET AL, "EXPRESSION OF E-CADHERIN ADHESION MOLECULES IN HUMAN BREAST CANCER TISSUES AND ITS RELATIONSHIP TO METASTASIS", PAGES 1696-1701, SEE ENTIRE DOCUMENT.	1-13

[X] Further documents are listed in the continuation of Box C.   [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 NOVEMBER 1994	JAN 19 1995

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	EGGERTON CAMPBELL
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PROC. NATL. ACAD. SCI. USA, VOLUME 87, ISSUED SEPTEMBER 1990, TSUDA ET AL, "ALLELE LOSS ON CHROMOSOME 16 ASSOCIATED WITH PROGRESSION OF HUMAN HEPATOCELLULAR CARCINOMA", PAGES 6791-6794, SEE ENTIRE DOCUMENT.	1-13
Y	CANCER RESEARCH, VOLUME 53, ISSUED AUGUST 1993, LINDBLOM ET AL, "DELETIONS ON CHROMOSOME 16 IN PRIMARY FAMILIAL BREAST CARCINOMAS ARE ASSOCIATED WITH DEVELOPMENT OF DISTANT METASTASES", PAGES 3707-3711, SEE ENTIRE DOCUMENT.	1-13
Y	CANCER RESEARCH, VOLUME 51, ISSUED NOVEMBER 1991, SATO ET AL, "ACCUMULATION OF GENETIC ALTERATIONS AND PROGRESSION OF PRIMARY BREAST CANCER",PAGES 5794-5799, SEE ENTIRE DOCUMENT.	1-20
Y	TECHNIQUE-A JOURNAL OF METHODS IN CELL AND MOLECULAR BIOLOGY, VOLUME 2, NO.3, ISSUED JUNE 1990, POLLARD-KNIGHT, "CURRENT METHODS IN NONRADIOACTIVE NUCLEIC ACID LABELING AND DETECTION", PAGES 113-132, SEE ENTIRE DOCUMENT.	1-23
Y	US, A, 4,996,143 (HELLER ET AL.) 26 FEBRUARY 1991, SEE ENTIRE DOCUMENT.	1-20